Regional Data and GIS Representation:
Methods, Approaches & Issues

Scoping Workshop for
GIS/Regionalization for EERE Models

July 15, 2004

NREL GIS Team
Liz Brady Sabeff, Ray George,
Donna Heimiller, Anelia Milbrandt

Our Goals

- Explore the use of Geographic Information System (GIS) applications and methods to improve model outputs from NEMS, MARKAL, and other EERE models
- Prioritize GIS/Regionalization development options for highest impact
- Support development of draft scoping paper for project

Objectives

- What types of data can easily be represented in GIS for renewable energy, census, and infrastructure
- Data improvements and existing GIS modeling efforts at NREL
- Discuss implications of scale & accuracy
- Definition of regions and data variation
- Scalability (spatial/temporal aggregation vs. computing time/sensitivity analysis/variability)
- Forecasting
- Data prep vs. modeling platform
- Issues around regionalization and modelers needs (parking lot for next day)

What Types of Data can Easily be Represented in GIS for Renewable Energy, Census, and Infrastructure?

- Several federal agencies have GIS data available in public domain (USGS, DOT, EPA, Census Bureau, BLM, …)
- State and local governments often have GIS data; but can be secondary and not well documented
- Vendors sell specialized or value-added data products related to energy, markets, weather, demographics, etc which may be GIS format.

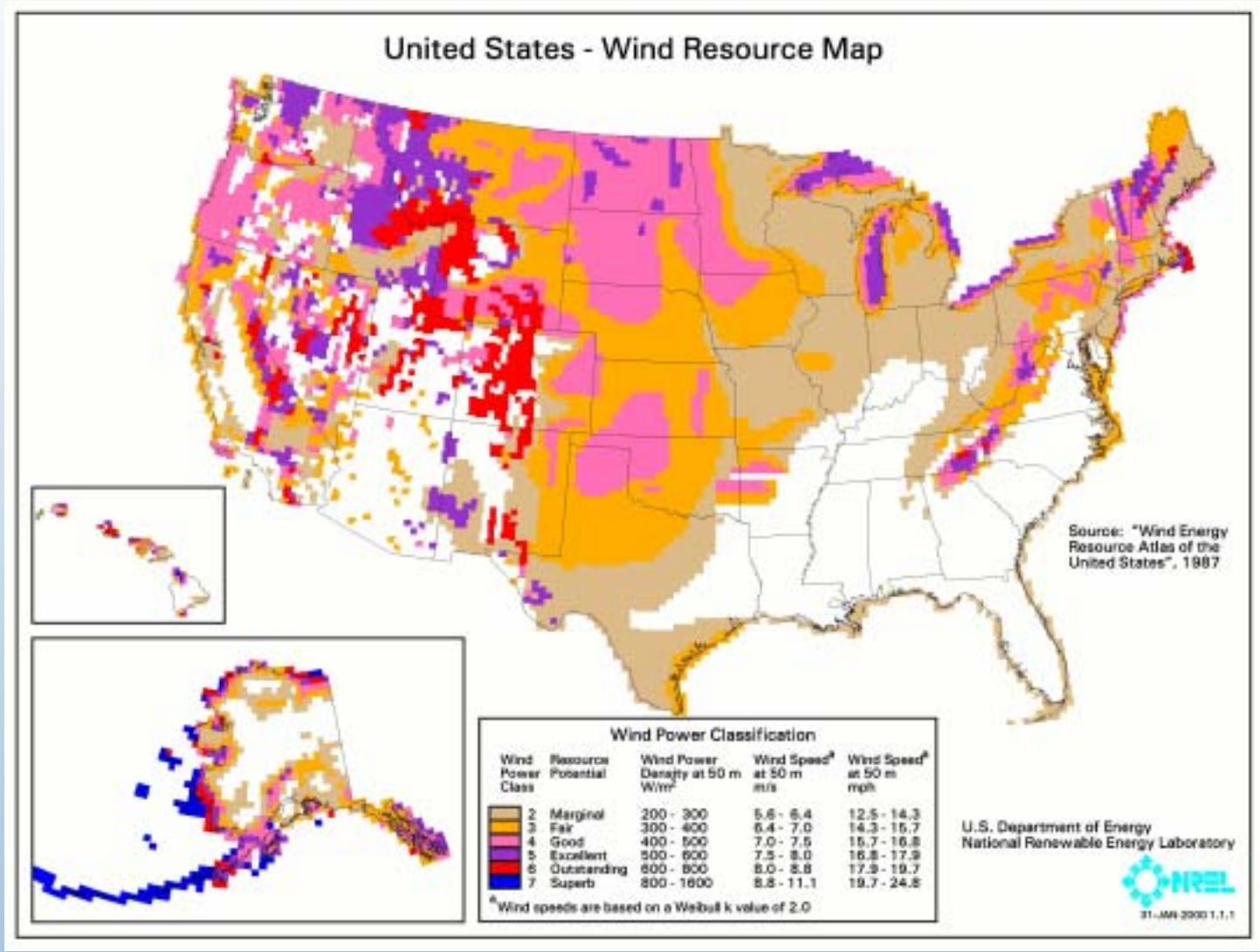
Databases Used with NREL GIS

- Renewable energy resources
- Utility industry information
  - Transmission (Electric and Gas), electricity rates
- Meteorological and climatological databases
  - Measured, modeled and derived
- Environmental
- Socio-Political
  - Demographic, administrative, federal lands
- Earth Science
  - Hydrography, elevation
  - Land cover, land use
  - Satellite imagery

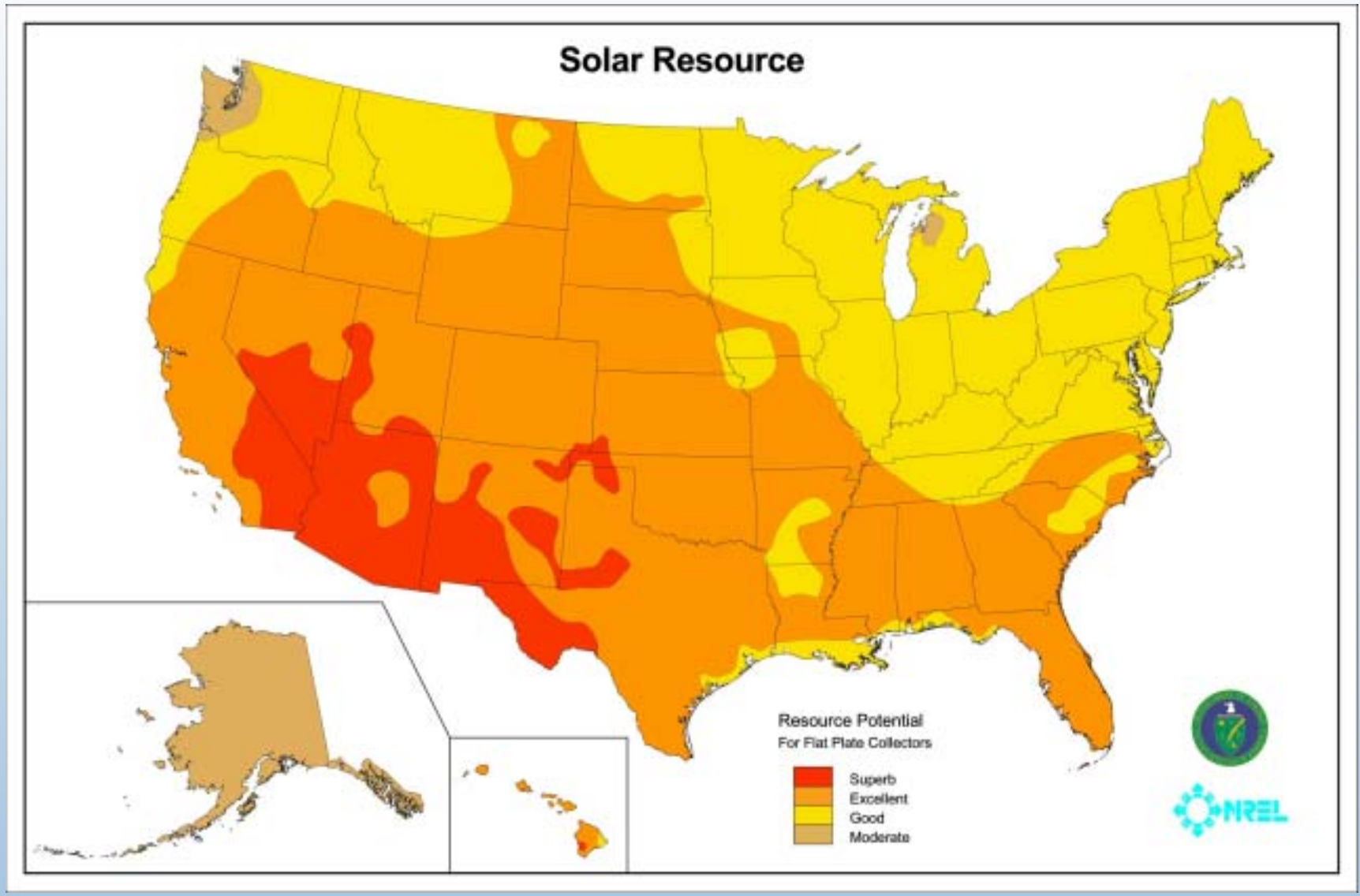
Renewable Energy Resources

- Wind – Modeled estimates
  - 1km to 200m (29 states), 25km 1987 for others. Upper air, surface and modeled meteorological data, satellite, and digital elevation models
- Solar – Modeled estimates
  - 40-km, 7-yr period, satellite and ground
  - 10-km for Southwest, 5-yr period, satellite
- Biomass – Statistical based estimates
  - NREL by county, 2004, data sources: USDA, USFS, ORNL, EPA, State Governments
  - ORNL by county,1999, includes economics
- Geothermal – Modeled estimates
  - Data from INEEL, SMU
- Hydrogen – under consideration/study

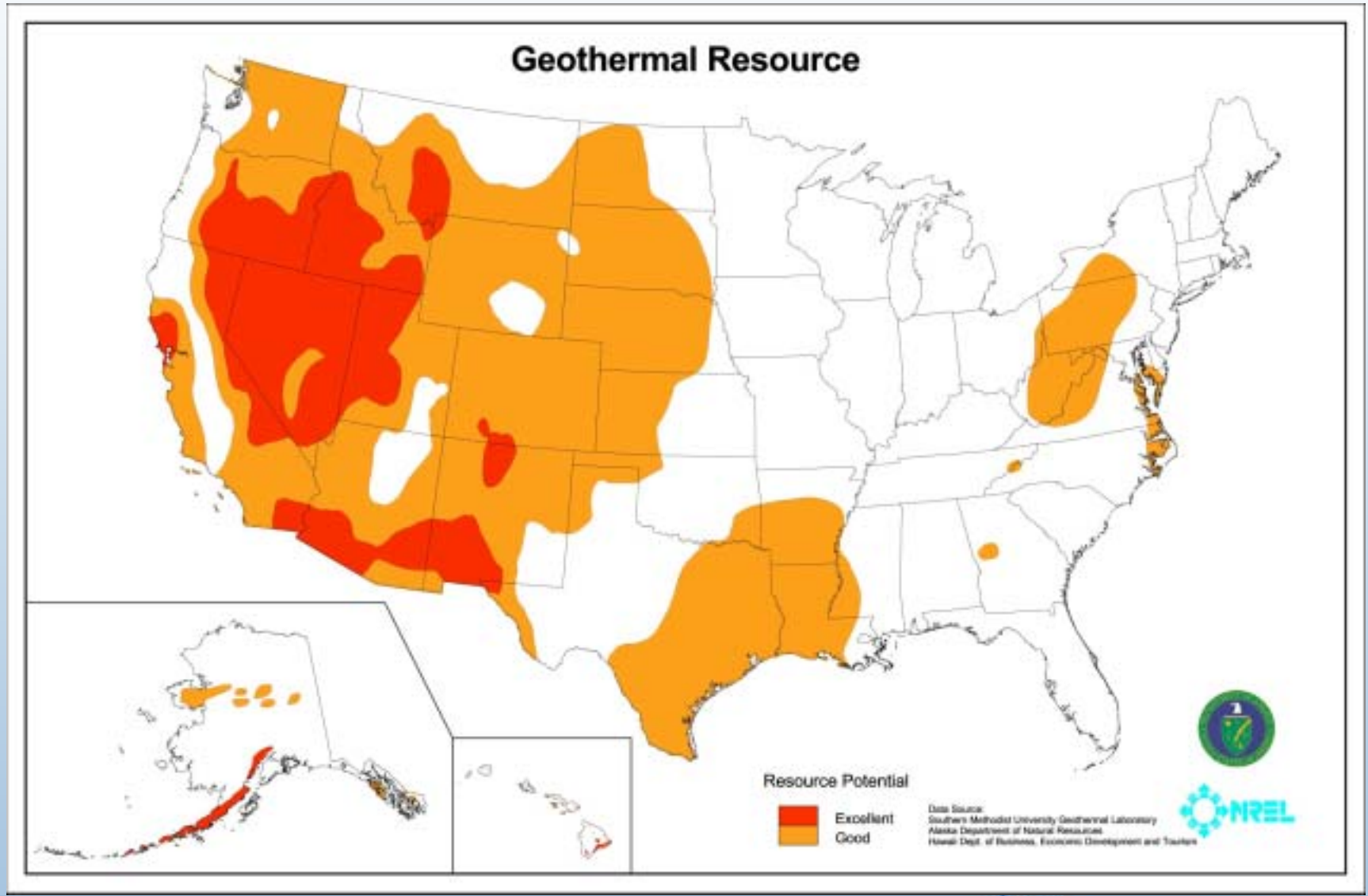
Renewable Energy Resources – Wind



Renewable Energy Resources – Solar



Renewable Energy Resources – Geothermal

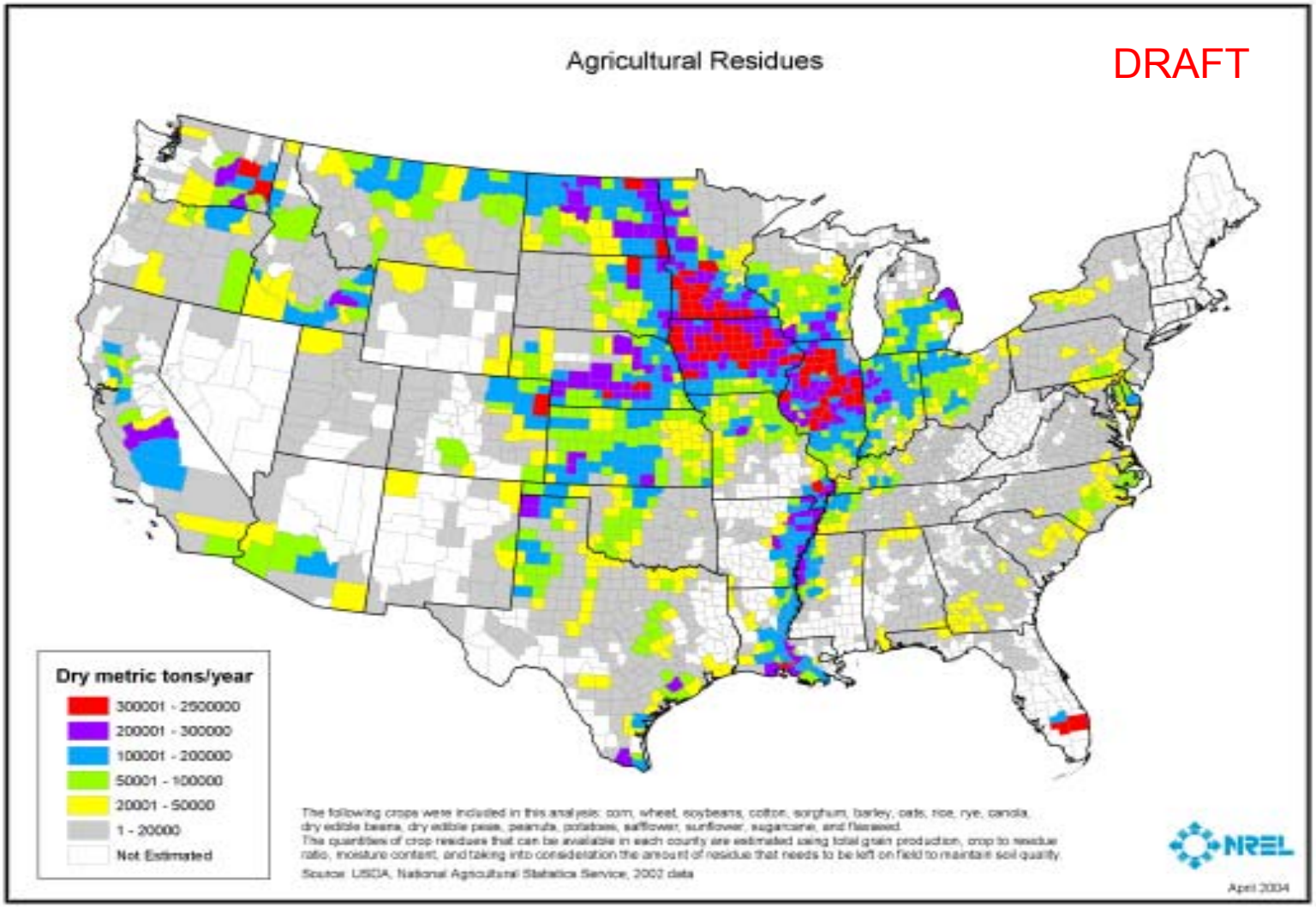


Renewable Energy Resources – Biomass Assessment

Statistical based estimates of the following biomass feedstock categories:

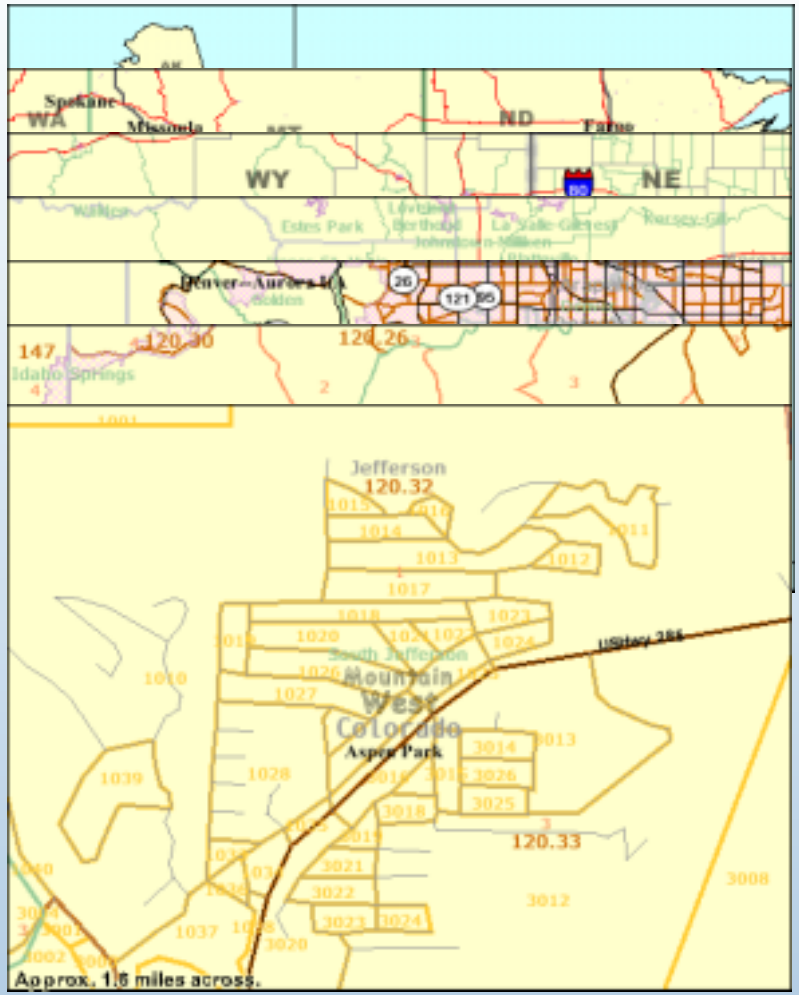
- Agricultural residues (crops residues)
- Livestock manure
- Landfill gas
- Energy crops
- Forest residues
- Primary Mill residue

Renewable Energy Resources – Biomass Assessment



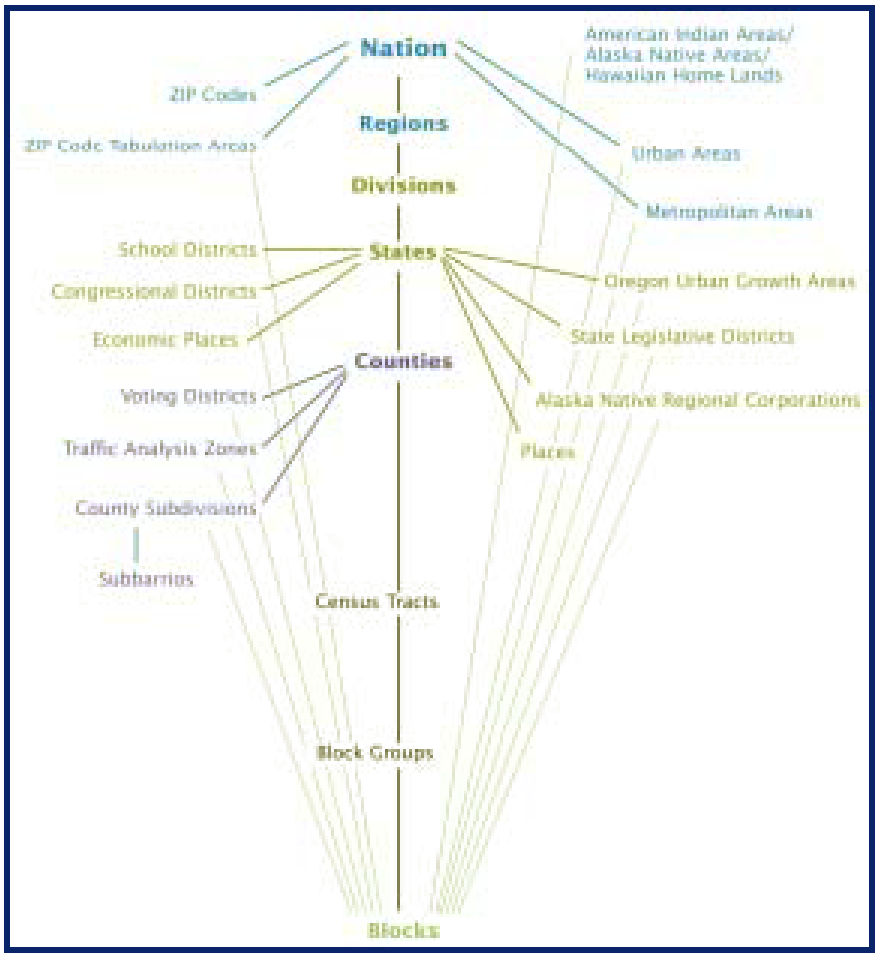
Census Bureau - GIS Data

- TIGER®
  - Digital database of geographic features, such as roads, railroads, rivers, lakes, legal boundaries, census statistical boundaries, etc. covering the entire United States. TIGER files do not contain demographic data, but can be "linked" to many of the other census data products
- Boundary Files
  - Generalized extract from TIGER in ESRI format for census areas, states, counties, school districts, tribal areas, congressional districts, voting districts, metropolitan areas



Census Bureau - GIS Data

Geographic Hierarchy



P15. HOUSEHOLDS [1] - Universe: Households				
National	Regional			
United States	northeast	midwest	south	west
105,480,101	20,285,622	24,734,532	38,015,214	22,444,733
State (northeast)				
Pennsylvania	New York	Connecticut	Maine	New Jersey
4,777,003	7,056,860	1,301,670	518,200	3,064,645
Counties (Colorado - sample)				
Adams	Arapahoe	Boulder	Denver	Jefferson
128,156	190,909	114,680	239,235	206,067
Jefferson County, CO CCD				
Central Jefferson CCD	Golden CCD	Northeast Jefferson CCD	South Jefferson CCD	
9,161	12,001	176,719	8,186	
Adams County, CO Census Tracts				
Tract 78	Tract 79	Tract 80	Tract 81	
2,714	1,944	1,877	370	

Census Bureau – Tabular Data

- People
  - Comprehensive population, sex, age, ethnicity & race, tenure. Sampled marriage, education level, languages spoken, occupation, industry, journey to work, income, home value, bedroom/bathroom count, home age, costs for utilities, taxes, mortgage, fuel, and insurance. Also migration patterns, census projections, historical census data,
- Business
  - Special EEO tabulation (gender, race, ethnicity, education, age, industry and earnings for nearly 500 occupations)
  - Economic Census reports number of establishments, revenue, payroll, and employees by North American Industry Classification System code
  - Business ownership characteristics
  - Government employment, revenue, expenditure
  - E-commerce statistics
  - Foreign Trade statistics

Population & Housing Examples

- Total population
- Urban and rural population
- Population by age, sex, and race
- Median age by sex
- Number of families
- Average family size
- Number of households
- Number of housing units
- Population in households
- Average household size
- Household size, type and presence of own children
- Household type by age of householder
- Household type by household size
- Tenure
- Occupancy status of housing

Infrastructure Data

- Roads, rail, waterways – ESRI, USDOT
- Electric Transmission & Natural Gas Pipelines - licensed from Platts
- Alternative fuel fueling stations
- Hydrogen production facilities
- Industrial facilities – USEPA, Platts
- EPAct Fleets
- Federal Facilities

Improvements in Data and Technology

Advances in geographic science have reduced cost and increased availability of information and applications to support geographic based analysis

- Evolution of Underlaying Geographic Sciences
  - Data Collection (remote sensing, LBS, GPS, surveys)
  - Data Types (demographic, business, marketing)
  - Data Modeling
- Evolution of GIS technologies
  - Single User Workstation to Multi-User Distributed
  - Internet Dissemination of Apps and data
  - Application Diversification (routing, statistics, etc)
- Planned improvements to NREL resource assessment products will provide better characterization of the resource with addition of diurnal, time series, and seasonal data tables

Examples of GIS support to modeling at NREL

- Wind, solar and biomass resource assessments
- Available windy land input (NEMS and WinDS)
- WinDS
- Clean Power Estimator
- PVWatts

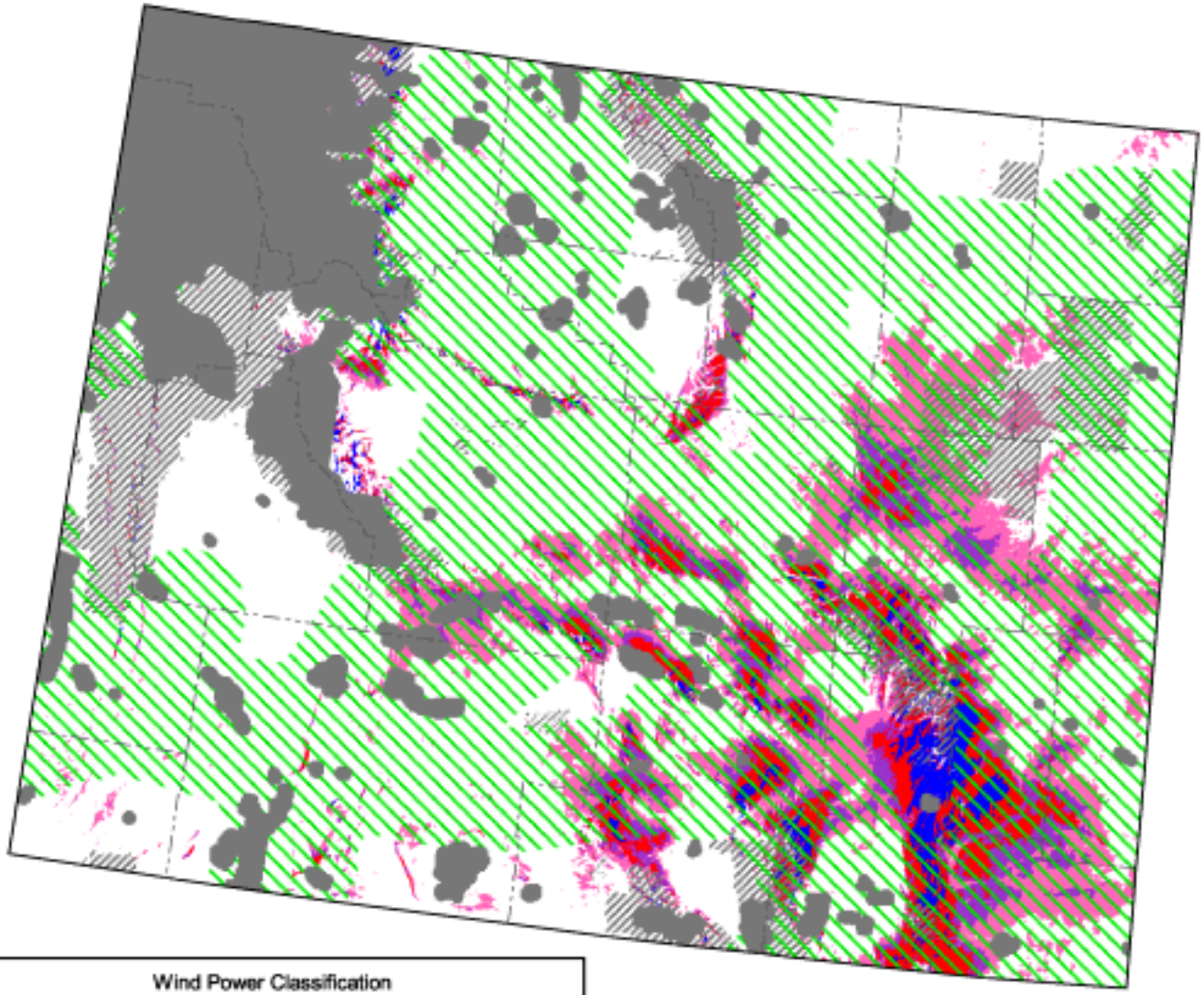
Update of Available Windy Land Input to NEMS

- NEMS used the 1987 U.S. wind resource data, with resource limited by distance to transmission lines as described in a 1995 report by Parsons et. al. (*Estimates of Wind Resource Land Area and Power Potential in Close Proximity to Existing Transmission Lines*)
- EIA approached NREL to update this input to NEMS to reflect the new high-resolution wind data sets currently available and more comprehensive transmission information available

Update of Available Windy Land Input to NEMS

- Analysis utilized the exclusion criteria developed by NREL with DOE guidance to quantify available windy land
- In addition, the resource is limited to within 20 miles of transmission lines
- Final input file has wind resource summarized by NEMS modeling region, power class and distance to transmission line (within 5, 10 and 20 miles)

NEMS Modeling Regions



Wyoming

Available Windy Land Within 20 Miles of Transmission 69 - 345 kV

The wind power resource data for this map was produced by TrueWind Solutions using the Mesomap system and historical weather data. It has been validated with available surface data by the National Renewable Energy Laboratory and wind energy meteorological consultants.

100% Excluded Areas
50% Excluded Areas
20 Mile Transmission Area

Wind Power Classification

Wind Power Class	Resource Potential	Wind Power Density at 50 m W/m ^{2}	Wind Speed [a] at 50 m m/s	Wind Speed [a] at 50 m mph
4	Good	400 - 500	7.0 - 7.5	15.7 - 16.8
5	Excellent	500 - 600	7.5 - 8.0	16.8 - 17.9
6	Outstanding	600 - 800	8.0 - 8.8	17.9 - 19.7
7	Superb	> 800	> 8.8	> 19.7

[a] Wind speeds are approximate and based on a Weibull k value of 2.0

40 0 40 80 120 160 Kilometers
25 0 25 50 75 100 Miles

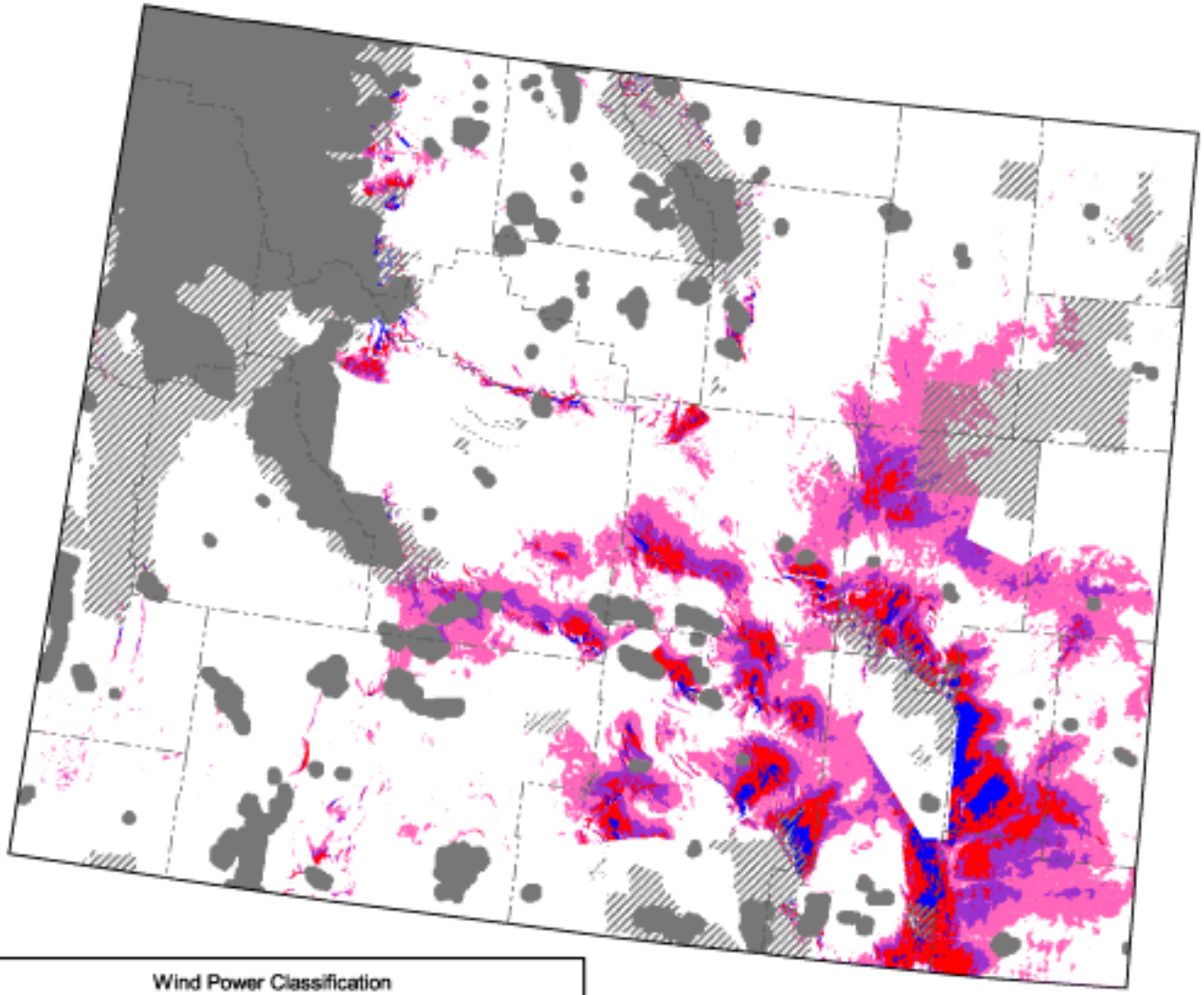
U.S. Department of Energy
National Renewable Energy Laboratory

NREL

14-JUL-2004 1.2.4

Wyoming

Available Windy Land Within 20 Miles of Transmission 69 - 345 kV



The wind power resource data for this map was produced by TrueWind Solutions using the Mesomap system and historical weather data. It has been validated with available surface data by the National Renewable Energy Laboratory and wind energy meteorological consultants.

100% Excluded Areas
50% Excluded Areas

Wind Power Classification

Wind Power Class	Resource Potential	Wind Power Density at 50 m W/m ^{2}	Wind Speed [a] at 50 m m/s	Wind Speed [a] at 50 m mph
4	Good	400 - 500	7.0 - 7.5	15.7 - 16.8
5	Excellent	500 - 600	7.5 - 8.0	16.8 - 17.9
6	Outstanding	600 - 800	8.0 - 8.8	17.9 - 19.7
7	Superb	> 800	> 8.8	> 19.7

[a] Wind speeds are approximate and based on a Weibull k value of 2.0

40 0 40 80 120 160 Kilometers
25 0 25 50 75 100 Miles

U.S. Department of Energy
National Renewable Energy Laboratory

NREL

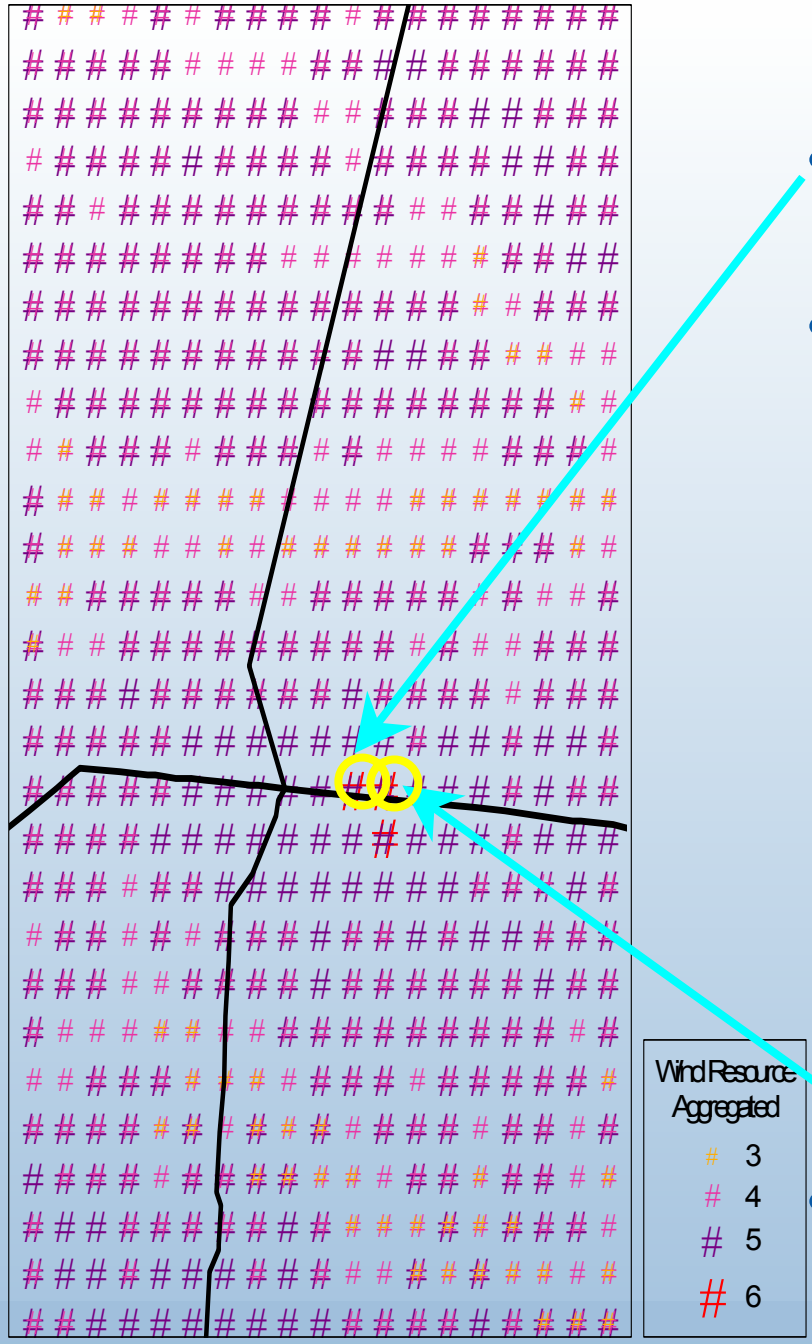
14-JUL-2004 1.2.4

WinDS GIS Processing

- GIS serves to prepare data inputs to a linear program
- Available windy land determined using exclusions based on the 1991 PNL windy lands assessment methodology (moderate land use exclusions)
- Assignment of wind resource to available transmission lines (used to generate regional wind supply curves)
- Demand, transmission capacity, load, current generation capacity and planned retirements allocated to the regions

WinDS GIS Processing: Transmission Assignment

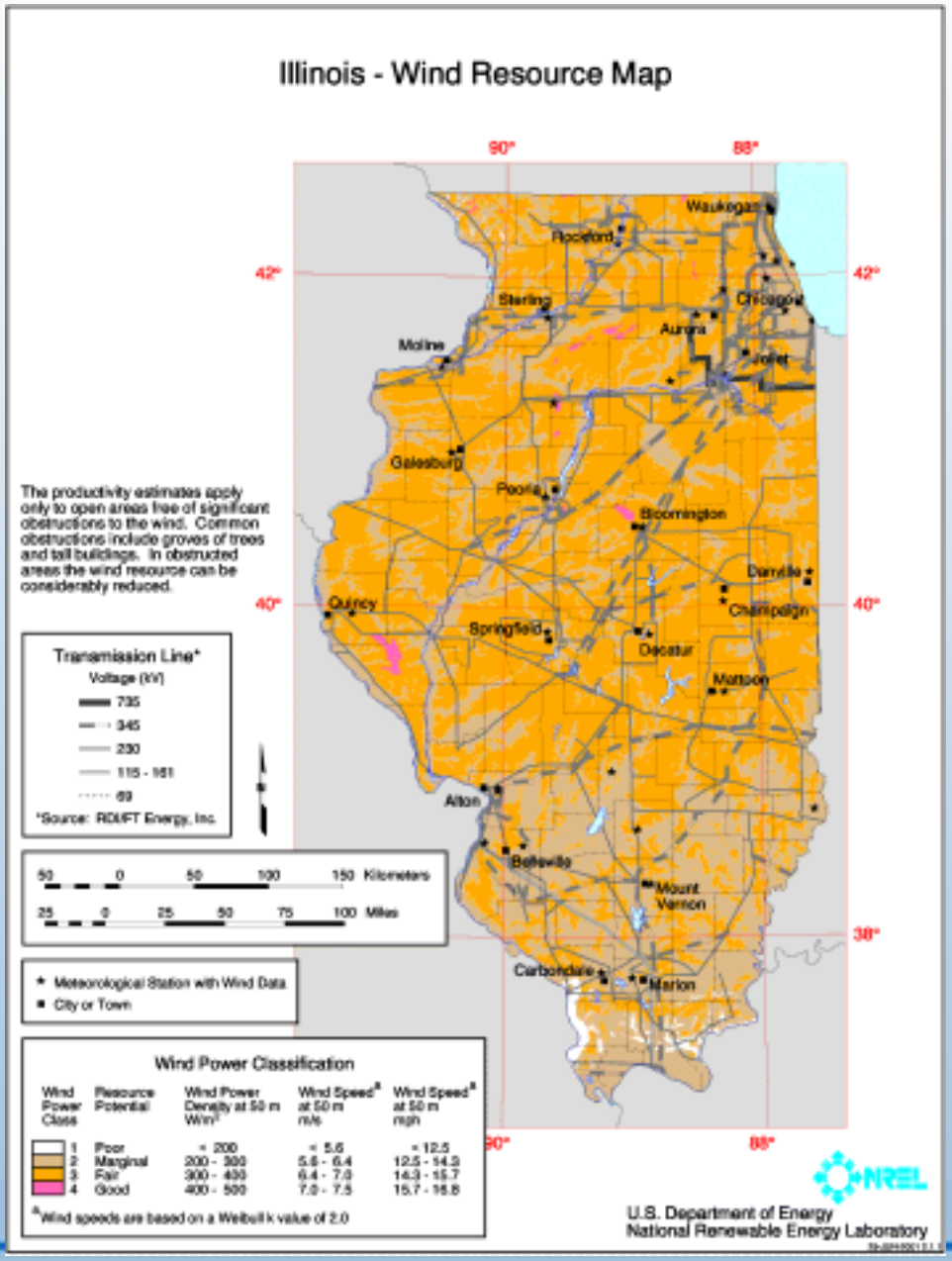
- Each wind cell assigned a cost based on the power class and distance to transmission
- Transmission line carrying capacity is based on voltage and length
- Wind cells are assigned to transmission lines based on cost, as long as they don't exceed the transmission lines available capacity
- Continues until all wind is assigned or all transmission capacity is used



- The list is sorted to identify the lowest cost wind cell
- The transmission line that it is currently assigned to is checked for available capacity
  - If there is capacity on the line, the point is assigned
  - In not, the next nearest line with enough capacity to take on that wind cell is found and assigned to that wind cell
- The next lowest cost cell is identified

WinDS GIS Processing: Transmission Assignment (cont.)

- The process repeats until all of the wind resource has been assigned or all of the the available transmission capacity has been used
- Allows lower power classes onto lines as long as they have the next lowest cost (levelized cost of energy + distance component of cost function)
- All wind records track what transmission line they were assigned to, the cost at which they were assigned, and what region they are in (used to generate wind supply curves for each region)



Clean Power Estimator

- Starts with high resolution annual wind power map
- Include analysis of daily and monthly wind speed patterns for meteorological stations in the surrounding area
- Need to be able to relate stations with good meteorological profiles to the grid cells

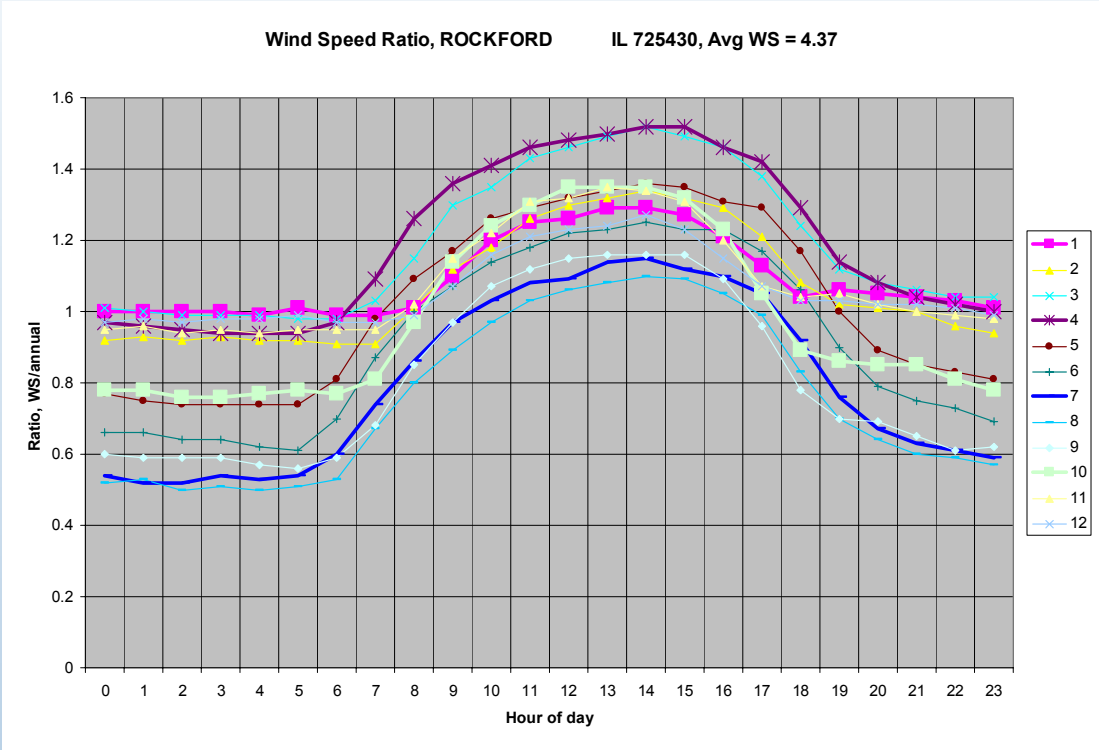
Clean Power Estimator for Wind

For each of 12 weather stations with ~ 10 m. winds

We graphed WS/WSann vs. hour and month

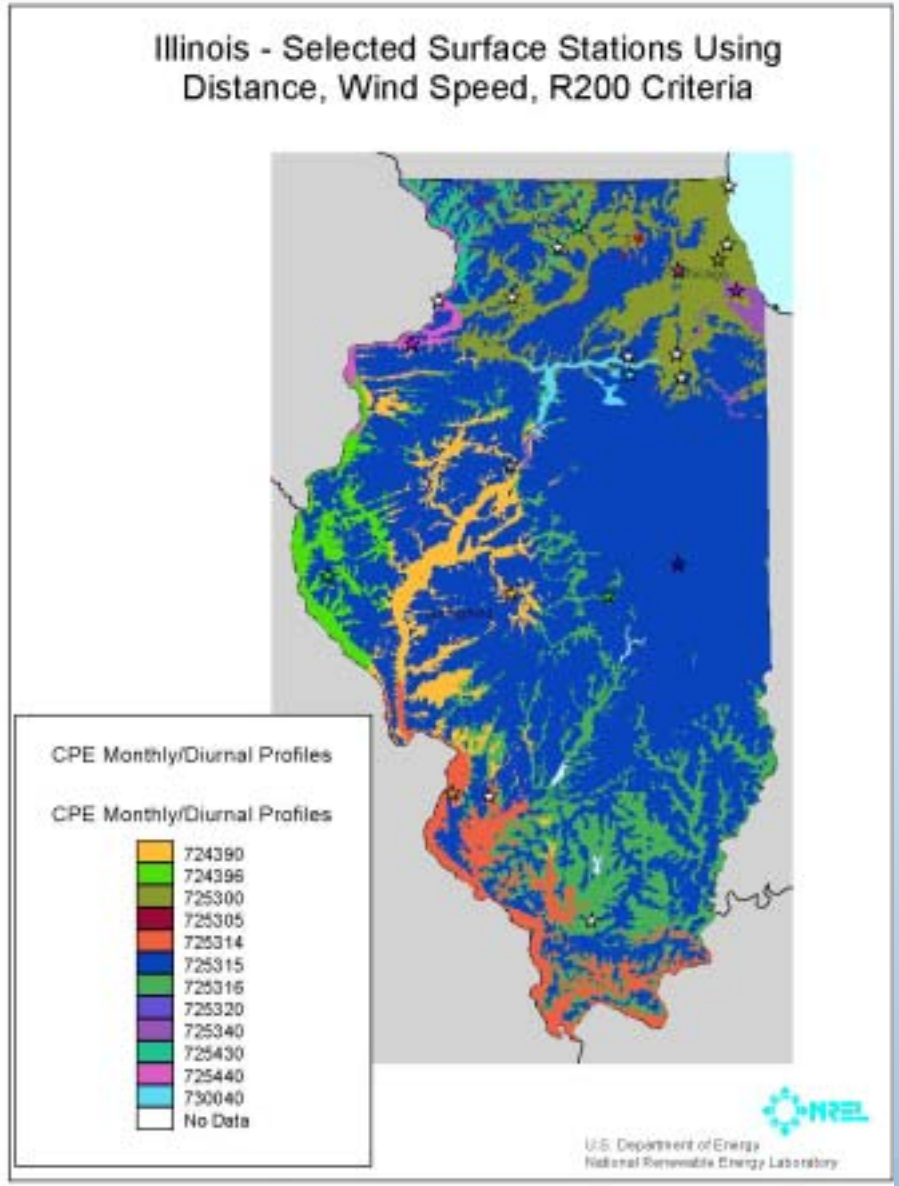
WS higher in the day, level at night.

Most stations poorly exposed.

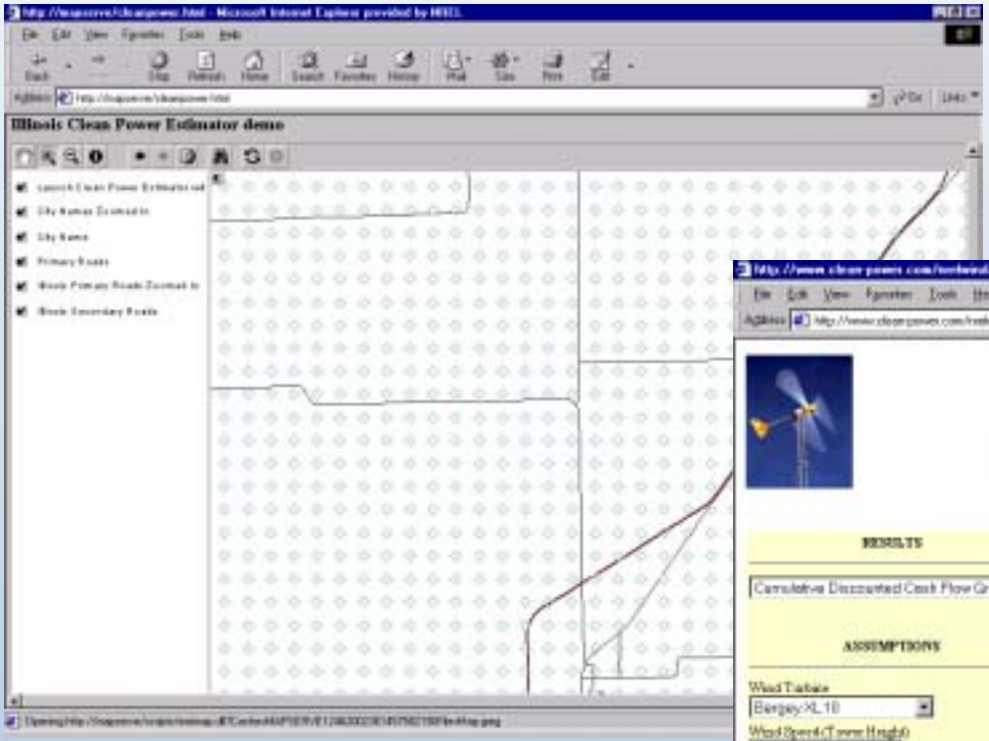


Chosen Station Profiles

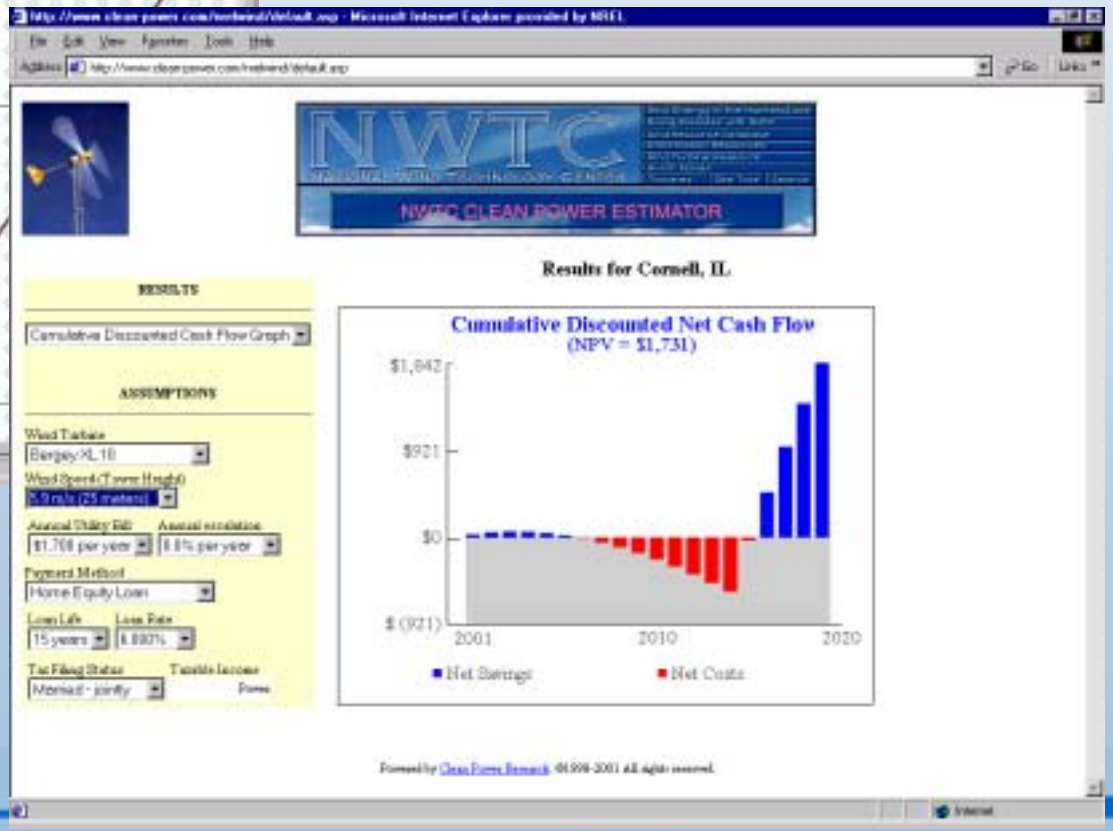
- Stations are assigned to each grid point by comparing the wind and terrain at the station with the wind and terrain at the grid point
- Choose the station that minimizes these differences:
  - Distance to station
  - R200 elevation
  - Wind speed at 10 m (30 m wind adjusted to 10 m using 1/7th Power Law)



Clean Power Estimator for Wind



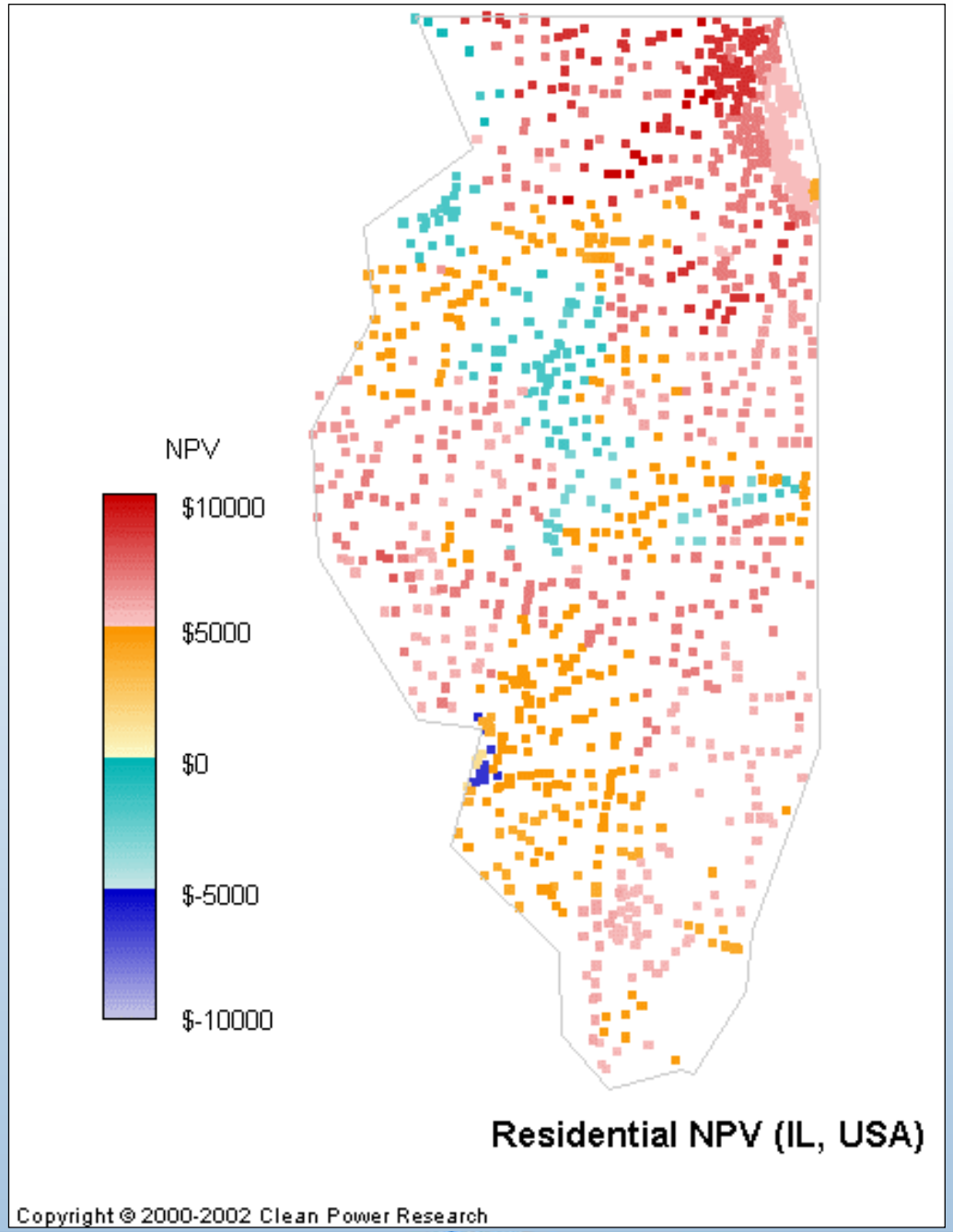
...Web point and click map server GUI, EQUALS ...



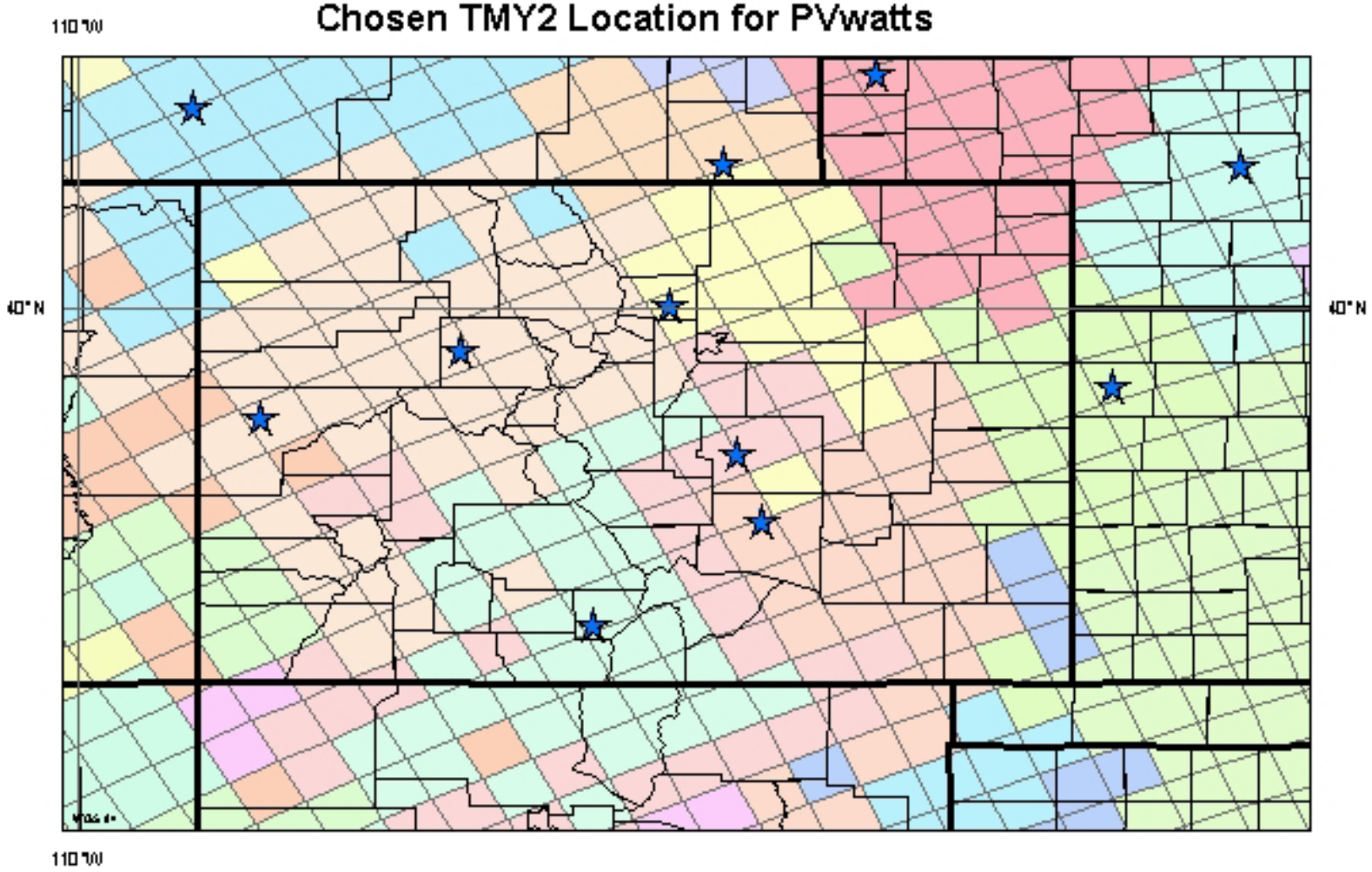
...Data for accurate estimation of wind turbine value under any utility rate scheme.

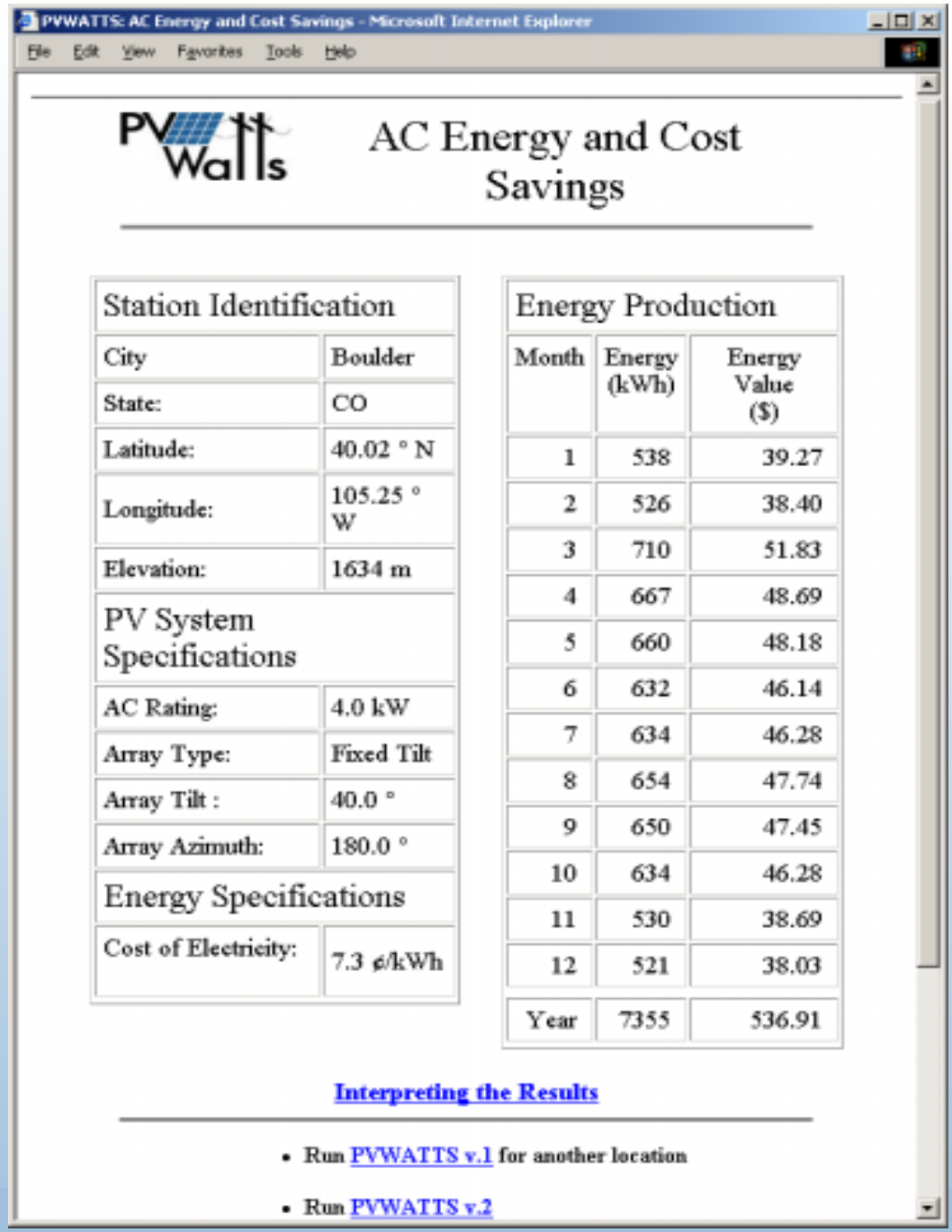
Net Present Value, Small Wind

- Values calculated using the Clean Power Estimator
- Each map point represents a city or location where the utility is known



Spatial Selection of Time Series

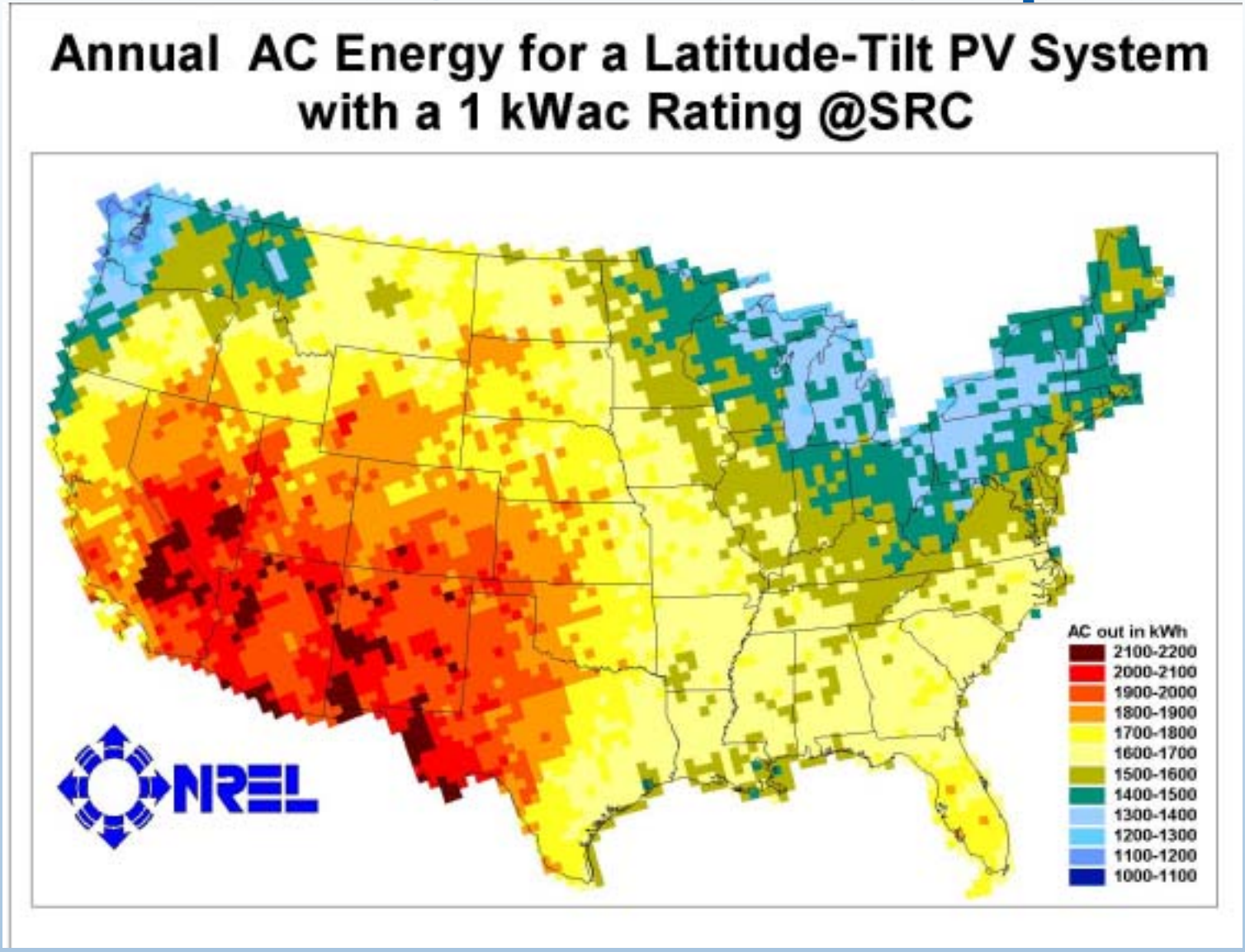




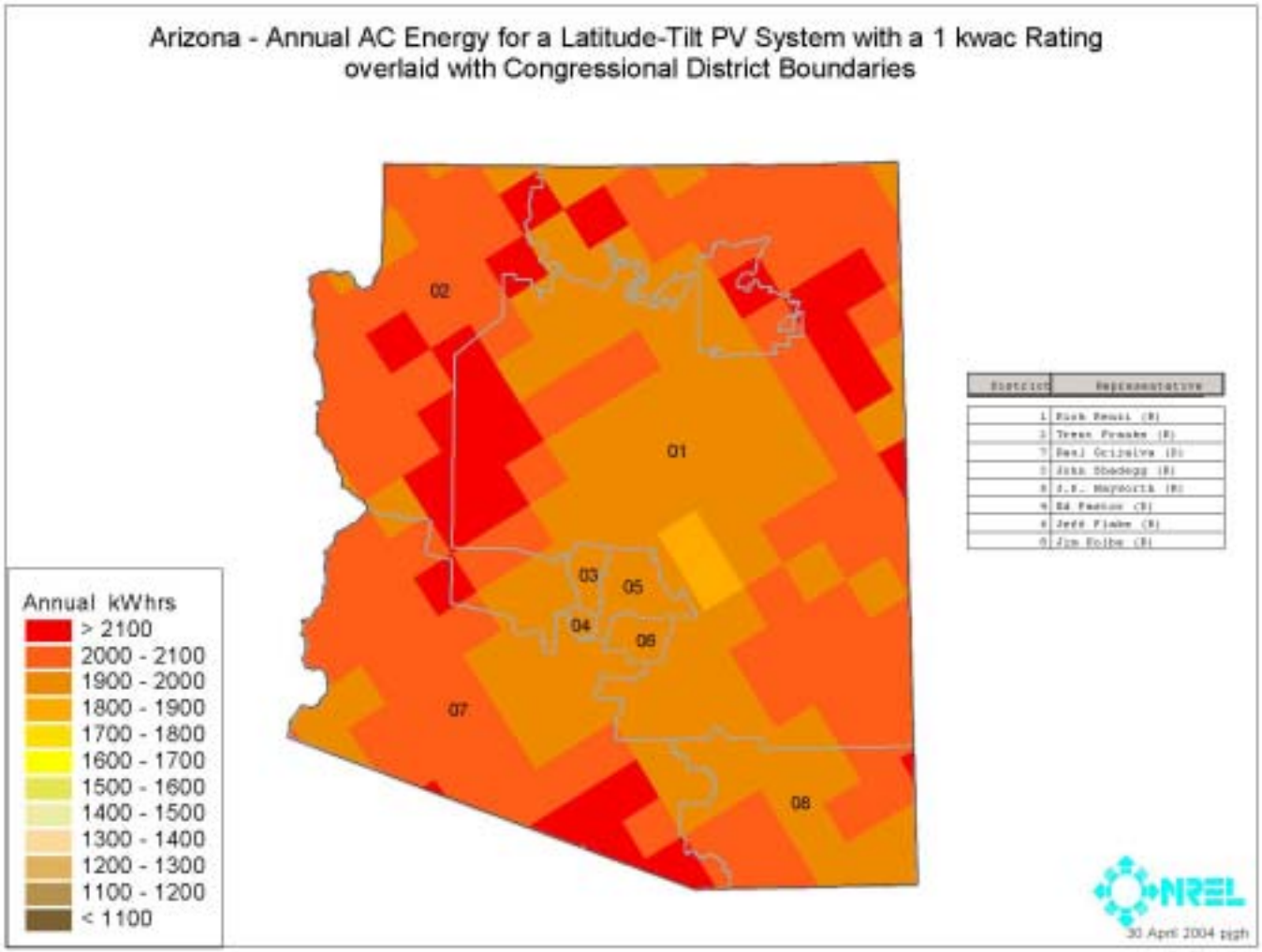
PV Watts

- 1 Year Hourly TMY Simulation
- Any Solar Collector Type and Orientation

Realistic Solar Model Outputs



Pvwatts Batch Output Spatial Coverage



Scale & Accuracy

- Consider how the data will be used and it's importance to the application
  - Taxing authority wants wetland boundary, environmental authority wants detailed wetland characteristics
  - City as a point for national level analysis, city as an area for local level analysis
- Don't use more detail than necessary
- Vintage of data will affect accuracy of analysis
- Source scale constrains level of detail in GIS dataset
- Location precision depends on scale, and a feature within a GIS represents the "probable" location

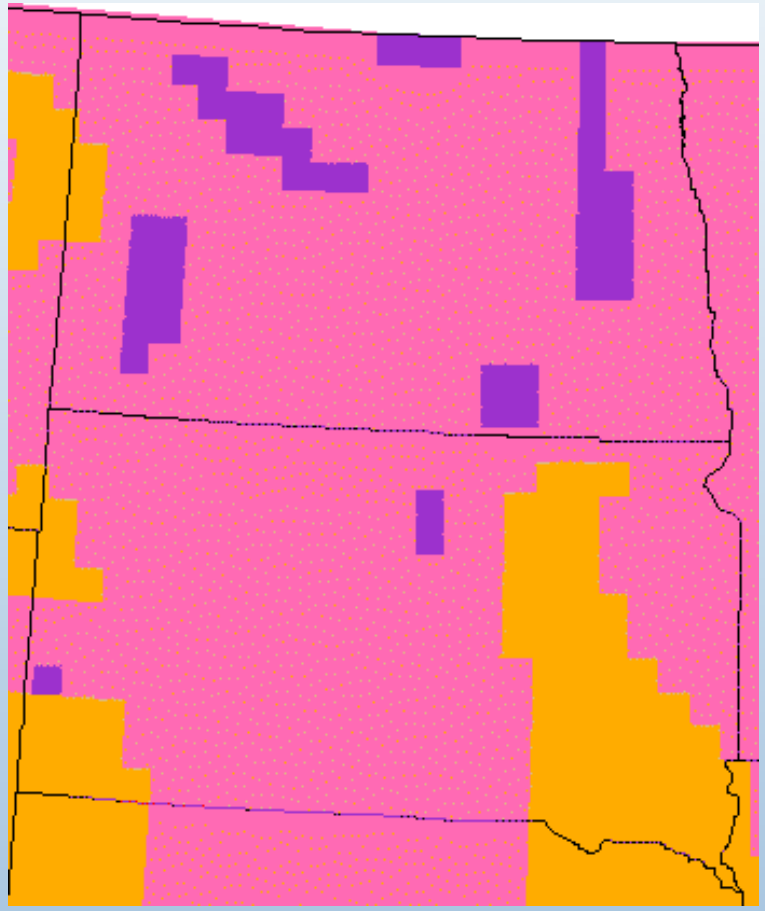
Data Quality Properties

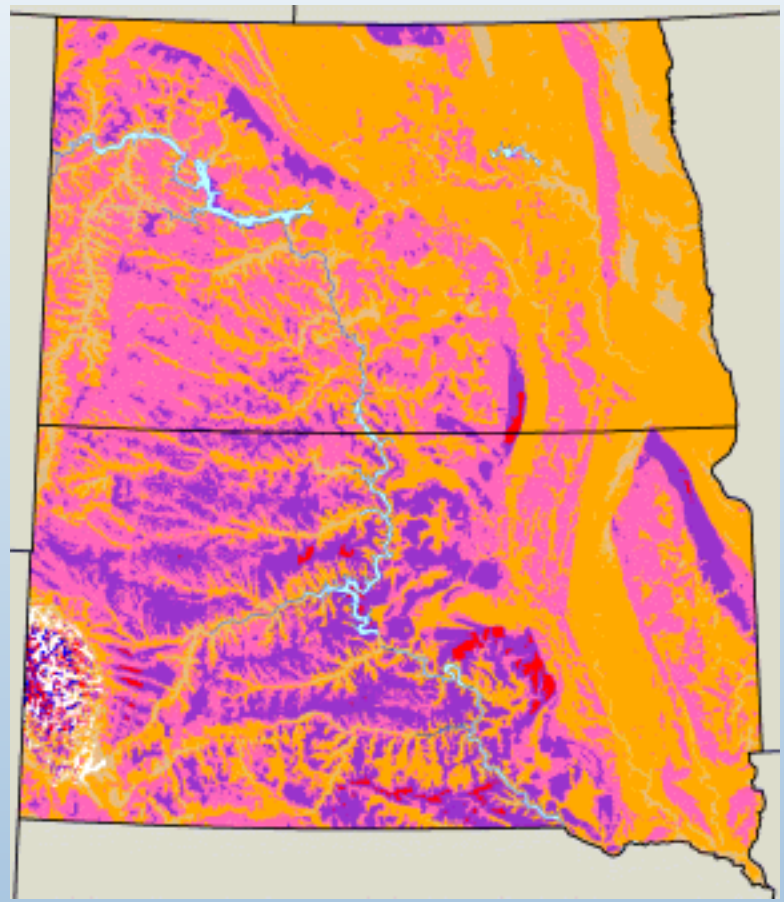
- Positional accuracy
- Attribute accuracy – are attributes correct?
- Conceptual accuracy – is data classified for analysis required?
- Logical accuracy and relevence – does this make sense?
- Completeness

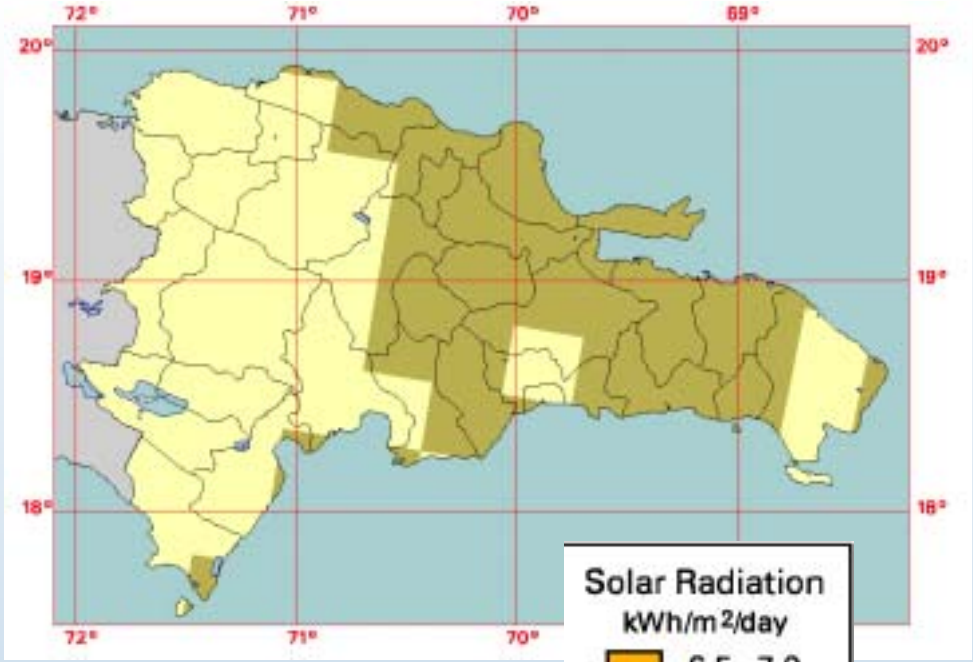
Error in GIS Databases

- Can be transferred to the GIS database because of errors in the original source (source errors)
- Can be added during data capture and storage (processing errors)
- Can occur when data are extracted from the computer
- Can arise when the various layers of data are combined in an analytical exercise
- Can occur when tabular data are linked to spatial data (classification errors)

1987 U.S. Wind Atlas Map vs. 2000 High-Resolution (1-km^2) Wind Map of North and South Dakota

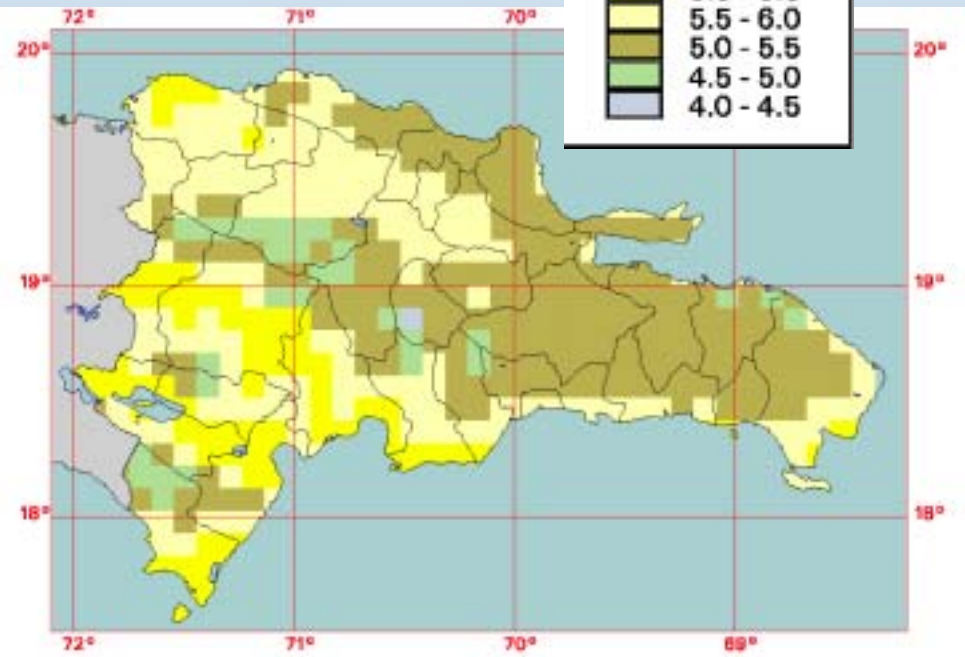






Estimates of annual average daily total radiation are modeled using inputs derived from satellite and/or surface observations of cloud cover, aerosol optical depth, precipitable water vapor, albedo, atmospheric pressure and ozone. Grid resolution is 40 km.

Global Horizontal Solar Resource



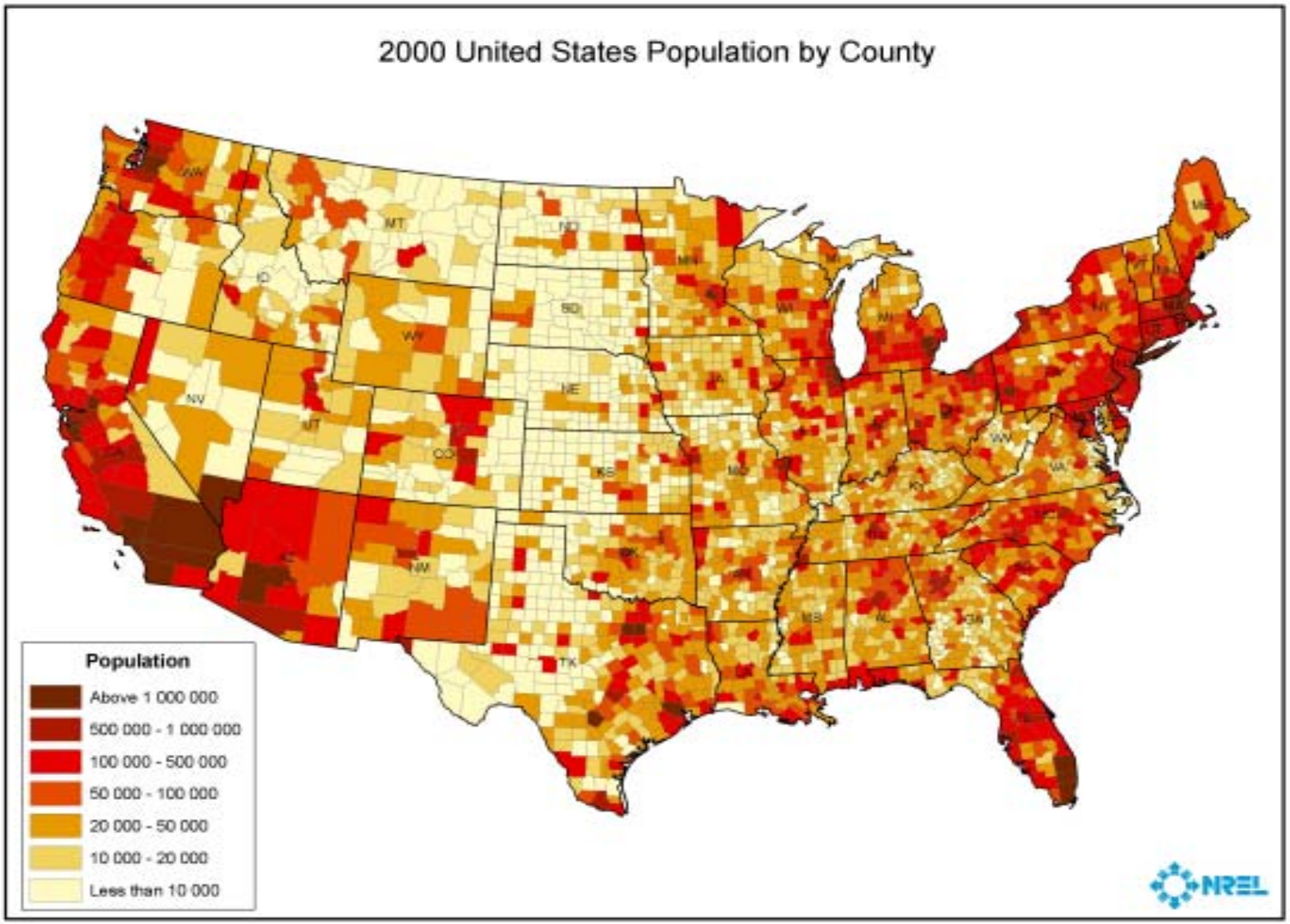
Comparison of NREL CSR Model and Perez' Satellite Model

Estimates of 1999 annual average daily total radiation are modeled using the visible channel of meteorological geostationary satellite imagery. Grid resolution is 10 km.

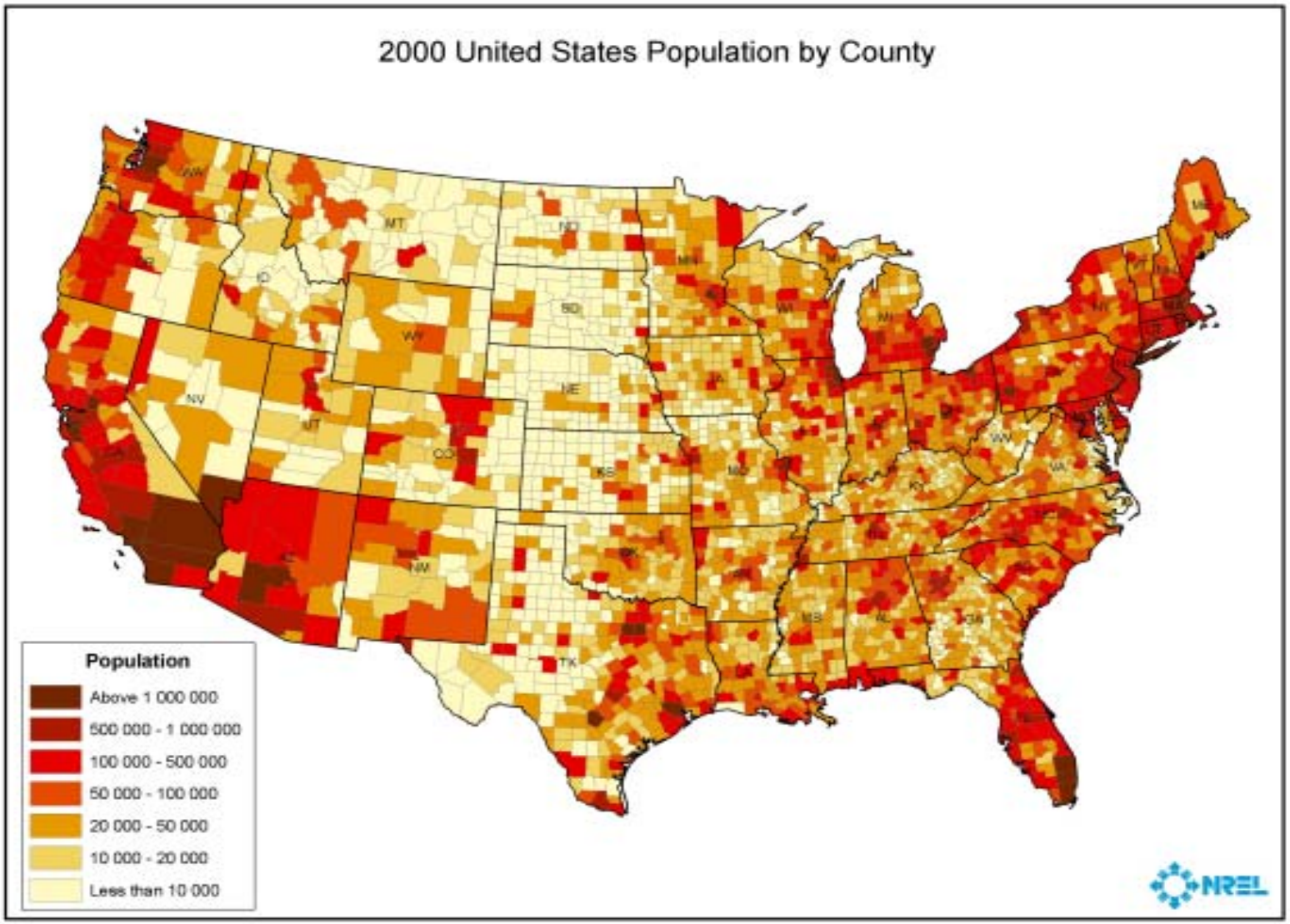
Regional Representation

- Different agencies use different regions, so combining data from these different sources can present challenges
- Region definition will vary based on need (political, economic, geological, environmental, etc)

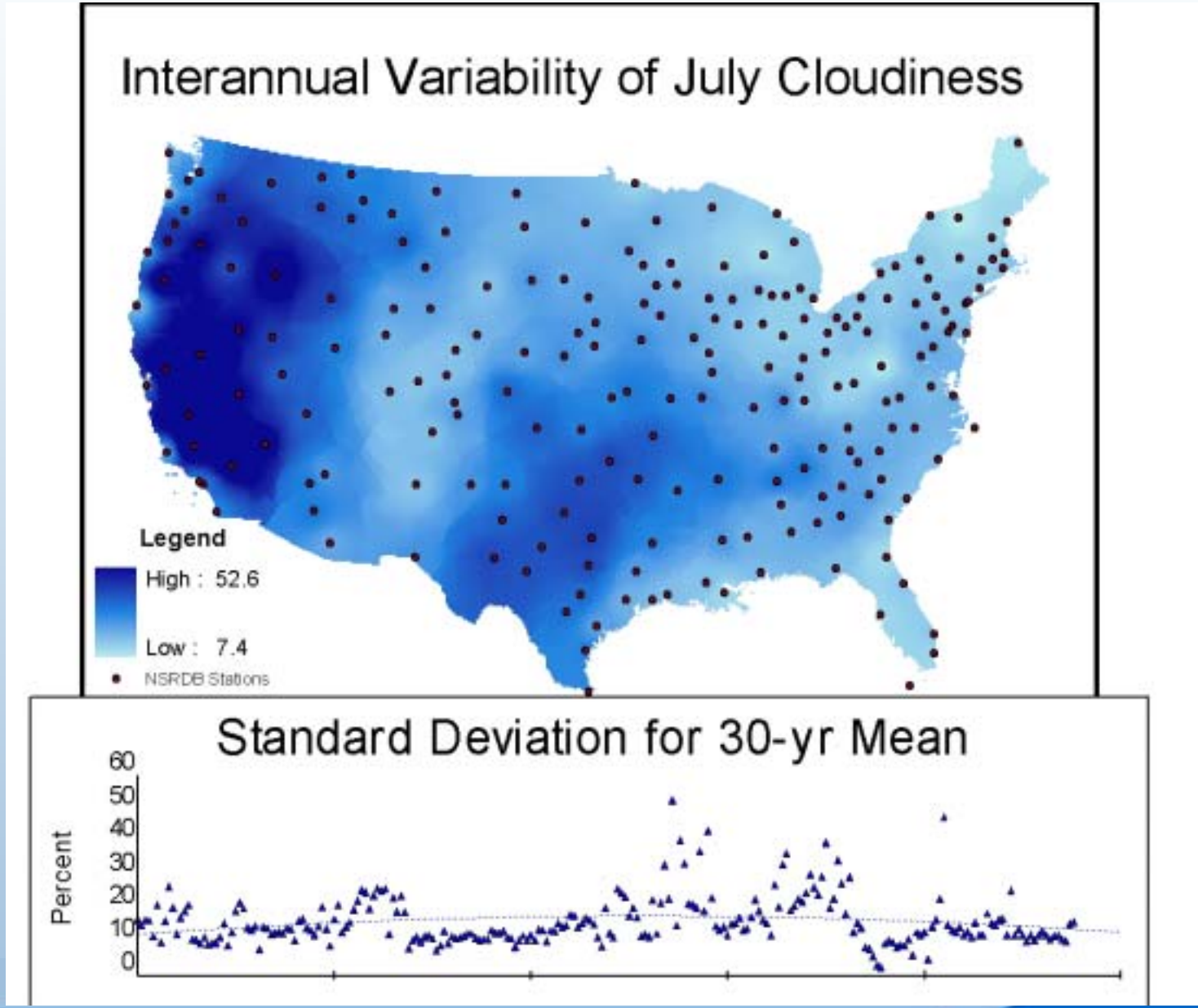
Geographic Variation Caused by Scale



Data Variation Caused by Classification



Geographic Variation Inherent to the Data



Scalability

- Transition from one model to another for differing region definitions
- Finer resolution aggregated up to some region, or coarser resolution disaggregated down to region
- Regional requirements for each model

Using GIS with Forecasting

- This is modeler's decision
- Depends on data requirements, computation time, complexity of model
- Look at modular approach
- Version control, documentation
- Model Builder

Using GIS for Data Prep Vs. Modeling

- Use GIS to create inputs to model (WinDS and CSR)
- Use GIS for model environment (WRAMS)
- Consider required inputs for model
- Consider spatial accuracy requirements
- Consider model environment (does model exist now?)

Issues

- Need to carefully consider data requirements
- Need to carefully evaluate suitability of data for desired model application
- Modeler's requirements must be clearly defined & communicated to GIS Analyst

GIS based regionalization should provide real value to improve model outputs